

# Differences in Multitask Resource Reallocation after Change in Task Values

Authors :

Nadine MATTON, ENAC, University of Toulouse and CLLE, University of Toulouse,  
CNRS, UT2J, France

Pierre PAUBEL, CLLE, University of Toulouse, CNRS, UT2J, France

Julien CEGARRA, Laboratoire Sciences de la Cognition, Technologie, Ergonomie (SCoTE),  
Université de Toulouse, INU Champollion, ALBI, France

Eric RAUFASTE, CLLE, University of Toulouse, CNRS, UT2J, France

Manuscript category: Research article

## **Abstract**

**Objective:** The objective was to characterize multitask resource reallocation strategies when managing subtasks with various assigned values.

**Background:** When solving a resource conflict in multitasking, Salvucci and Taatgen (2008; 2011) predict a globally rational strategy will be followed that favors the most urgent subtask and optimizes global performance. However, Katidioti and Taatgen (2014) identified a locally rational strategy that optimizes only a subcomponent of the whole task leading to detrimental consequences on global performance. Moreover, the question remains open whether expertise would have an impact on the choice of the strategy.

**Method:** We adopted a multitask environment used for pilot selection with a change in emphasis on two out of four subtasks while all subtasks had to be maintained over a minimum performance. A laboratory eye-tracking study contrasted 20 recently selected pilot students, considered as experienced with this task, and 15 university students considered as novices.

**Results:** When two subtasks were emphasized, novices focused their resources particularly on one high-value subtask and failed to prevent both low-value subtasks falling below minimum performance. On the contrary, experienced people delayed the processing of one low-value subtask but managed to optimize global performance.

**Conclusion:** In a multitasking environment where some subtasks are emphasized, novices follow a locally rational strategy whereas experienced participants follow a globally rational strategy.

**Application:** During complex training, trainees are only able to adjust their resource allocation strategy to subtask emphasis changes once they are familiar with the multitasking environment.

**Keywords:** Adaptability, Time sharing, Experience, Eye tracking, Attentional processes

Précis:

In a demanding multitasking environment, novices and experienced participants reacted differently when some subtasks were emphasized while others were underemphasized. Novices disregarded the two low-value subtasks whereas experienced participants still maximized their global performance.

## INTRODUCTION

Nowadays, many work environments require management of concurrent subtasks. As human ability to process more than one task at the same time is severely limited (e.g., Loukopoulos, Dismukes & Barshi, 2009), operators often have to set priorities to the different subtasks and process them in the corresponding order. However, the values of the subtasks are likely to vary across time, leading operators to also change their resource allocation strategy in order to adapt to the new set of subtask values. During training, a robust finding is that trainees confronted with a multitasking environment with variable priorities master the task faster than trainees who are only confronted with fixed priorities (see Gopher, 2007 for a review and Boot et al., 2010 for recent results). The added-value of the variable-priority approach is supposed to come from the opportunity to explore various strategies which would be beneficial for learning to manage the multitasking environment (e.g., Kramer, Larish & Strayer, 1995). In the variable-priority approach, participants are assigned to either experimental condition. Variable-priority learners are instructed to focus their attention on specific subcomponents of the task at different moments. On the contrary, full emphasis training learners are instructed that each subtask is equally important. However, the variable-priority paradigm is not devoted to understanding the attention reallocation strategy when switching from an equal-emphasis stage to a differential-emphasis stage. The current study

aimed at better understanding attention reallocation in a multitasking environment when some subtasks become temporarily more important than others. More specifically, we focused on differences in reallocation strategies between novices and experienced users of a given multitasking environment.

Several models have been established to describe multitasking behavior. For instance, the STOM (Strategic Task Overload Management) model makes predictions of multitasking performance of an overloaded operator who may decide to switch from an ongoing task to alternative tasks based on task attributes (Wickens, Gutzwiller & Santamaria, 2015 ; Wickens et al., 2016). Five task attributes have been defined in this model: Saliency of the task, effort of task switching, task priority that comes from the instructions, task interest, and task difficulty. Thus, if only task priority changes from one stage to another, the STOM model predicts that people will focus for a longer period on a high-value subtask than on a low-value one. However, this model does not specify what operators do with the low-value subtask. Indeed, in most multitasking environments, paying less attention to one subtask may have detrimental effects on global performance, so the reallocation of attention should also take into account the characteristics of the multitasking environment in order to optimize global performance. Another multitasking model, threaded cognition (Salvucci & Taatgen, 2008, 2011), specifies that in case of conflict of several subtask goals for the same cognitive resource (e.g., the procedural resource that initiates the production rules), the subtask goal *with the highest urgency* will be proceeded. Thus, when a subtask value is temporarily increased, one can suppose that its urgency will increase so that the subtask will be processed earlier. However, if the low-value subtask performance is so poor that it is a threat for the global performance, it will become urgent to process this low-value subtask. Therefore, threaded cognition is well suitable to describe optimal behavior, i.e. behavior that will lead to optimal global performance.

In contrast, Katidioti and Taatgen (2014) have highlighted that in high workload conditions, people may adopt a *locally rational* strategy (that optimizes only one sub-aspect of the task) rather than a *globally rational* strategy (that optimizes global performance). More specifically, in their study participants switched from a primary email task to a secondary chat task mainly when their resources were available (locally rational), although global performance thereby decreased (globally irrational). In the same vein, Fu and Gray (2006) have found that in a map-navigation task, participants had the tendency to follow a suboptimal strategy that led to a local optimum (by stopping environment exploration prematurely) rather than an optimal strategy that would lead to a global optimum (by using more information seeking actions and using a path that initially went diverged from the end station). Moreover, the switch from one task to another is known to result in slower response latencies and higher error rates (*switch cost*, Monsell, 2003). Indeed, task switching requires reconfiguring the procedural schema (Norman & Shallice, 1986) in order to select and implement appropriate actions given the dominant goal (Duncan, 1990). Thus, when the value of one subtask increases, one can suppose that there is some reluctance to switch to low-value subtasks. Analogously, the cognitive effort avoidance principle (e.g., Kahneman, 2011) or the minimal control principle (Taatgen, 2007) would predict that when some subtasks become less important, people will prefer a resource allocation strategy that focuses on processing the minimal number of subtasks. In particular, this principle predicts that the subtasks selected will be those of high value. To sum up, a change in subtask values may lead to two kinds of strategies, (1) a strategy that maximizes global performance or (2) a strategy that only maximizes some subcomponent(s) of the task. One can assume that depending on the experience people have of the multitasking environment, the preferred strategy may not be the same.

Indeed, many research findings showed that resource allocation strategies differ between experts and novices. In aviation, expert pilots are known to adapt their visual attention

allocation more flexibly to changing task demands than novice pilots (e.g., Bellenkes, Wickens & Kramer, 1997). Expert helicopter pilots also showed more effective visual allocation strategy than inexperienced pilots during a demanding mission (Robinski & Stein, 2013). In car driving, Crundall and Underwood (1998) highlighted that experienced drivers adjusted their scanning patterns to different types of roads whereas novices may be insensitive to the road type (measured through the variance of locations of fixations). Likewise Underwood, Chapman, Brocklehurst, Underwood and Crundall (2003) found differences in visual attention strategies between novice and experienced drivers. Experienced drivers showed greater sensitivity overall and novices demonstrated more stereotypical patterns.

### **Current Study**

The present paper examines the possibility that, in a multitasking environment, after subtask value changes, novices would follow a *locally rational* strategy that focuses more frequently on high-value subtasks rather than low-value subtasks, whereas experienced people would follow a *globally rational* strategy that maximizes the global performance in paying attention to each subtask when it becomes urgent. We used a multitasking environment in which the locally rational strategy leads to a poor global performance, thus enabling both strategies to be separated. In this environment participants had to manage four different subtasks in two conditions: Firstly the four subtasks were explicitly instructed as being equally important, and secondly two of the four subtasks were more important than the two others. All other task characteristics remained constant in both conditions. An important characteristic of the global performance calculation is that it is both a linear combination of each subtask performance but it also follows a disjunctive rule. This rule posits that global performance is set at zero as soon as one subtask performance is below a 5% threshold. After the change in subtask emphasis, a locally rational strategy would lead to neglect the two low-value subtasks and hence to produce a suboptimal performance. On the contrary, the globally

rational strategy would lead participants to prioritize subtask processing according to their assigned utilities. For instance, to maximize performance, high-value subtasks would be processed with high priority and low-value subtasks with low priority. However, a low-value subtask should be processed as soon as it becomes urgent in order to prevent detrimental effects on global performance. Moreover, as noted by Gopher, Weil and Siegel (1989, p. 154), in general when subtask emphasis is changed, “the task is not simplified (...) because in addition to the initial requirements, subjects must also monitor priorities”. Therefore, we hypothesize that overall performance would decrease after the change in subtask values for both groups. We also expect that the decrease in performance would be greater for novices, as they are supposed to have fewer available resources for dealing with this additional complexity than experienced participants.

### **General Method**

#### **Overview**

In the two empirical studies reported here, we used the same multitasking environment and performance measurements. Firstly, a preliminary large scale field study showed that performance decreased overall after the change in subtask values and that the decrease in performance was greater for first-takers than for more experienced participants. Secondly, an experimental eye-tracking study aimed at contrasting visual attention strategies in novices and experienced participants. Data of the preliminary study (behavioral performance measurements only) were collected during an actual pilot-training selection process, whereas the eye-tracking study was an experiment conducted in laboratory conditions. This research complied with the American Psychological Association Code of Ethics.

#### **Experimental setup**

The experiment consisted in a computerized task. Participants sat in front of a 19-inch and 1024x768 resolution computer screen and interacted through two joysticks and a keyboard. The procedure will be described more precisely for each study.

### **The Priority Management Task**

The Priority Management Task consisted in the simultaneous completion of four concurrent subtasks and was organized in six successive four-minute stages. During the first four stages, the subtasks were successively added (from only one subtask at the first stage to four subtasks at the fourth stage) in order to gradually familiarize the participants with the management of the four subtasks. Subtasks were appended in the same order for all participants: Monitoring, Tracking, Detecting and Calculating. Thus, at the end of the six stages participants were more familiar with the Monitoring than with the Calculating subtask. However, the purpose here was not to compare performance of subtasks. All the characteristics of the subtasks were exactly the same for each participant. An event requiring an action of the participant was triggered every 15s for all subtasks. Moreover, the four subtask events were synchronized, forcing the participant to make choices. The ‘monitoring task’ consisted in maintaining the levels of four gauges within an interval by using a first joystick. Every 15s one of the gauges deviated from its position at a speed of 10 to 70 pixels per second. To maintain the level of the gauge in the target interval, the participant had to press on several buttons in order to select the right gauge and then to adjust the gauge to the desired value with the joystick. The ‘tracking task’ consisted in keeping a cross positioned in a moving circle of 50 pixels diameter, in an area bordered by a large circle of 300 pixels diameter, through the second joystick. The circle moved every 15s at a speed of 3 to 12 pixels per second. The ‘detection task’ consisted in detecting the presence of three target letters (which varied from stage to stage) in a block of nine letters. Participants had to press one of nine keyboard keys as quickly as possible when a target letter appeared in the corresponding zone. A new block of letters was presented every 15s. The ‘mental calculation task’ consisted



of simple arithmetic problems (e.g., deducing a distance from given speed and time). The participants had to type the numeric answer as quickly as possible. Whether an answer had been given or not, a new problem was presented every 15s.

For each subtask, the instantaneous performance level was displayed through a corresponding gauge at the top center of the screen. In this paper we focused on stages S4 to S6, the only stages where participants had to manage the four subtasks simultaneously. The characteristics of the four subtasks remained the same across the three stages. The only difference across stages was the weighting attributed to each subtask. Indeed, during these stages, a global performance gauge on the right of the subtasks performance gauges displayed instantaneous global performance computed from the four subtasks (details in the performance measurements subsection):

- At stage S4-equal the four subtasks were defined as equally important (through explicit instructions and a reminder of the four relative weightings during the whole four-minute stage, see Figure 1, left panel).
- During stage S5-diff, participants were informed that two subtasks were more important than the two others and that they had to comply with these assigned values. However, they were also explicitly instructed not to totally neglect the low-value subtasks. The assigned values were again presented during the whole stage (see Figure 1, right panel). The choice had been made to put the emphasis on the two “less salient” tasks (letter detection and mental calculation), as they were supposed to capture less attention than the two other tasks (tracking and monitoring) which comprised moving targets (e.g., McLeod, Driver & Crisp, 1998). Therefore, during S5-diff, Monitoring and Tracking were “low-value” subtasks whereas Detecting and Calculating were “high-value” subtasks.
- During stage S6-equal, the four subtasks were equally important, as during S4-equal and all characteristics were identical to S4-equal. This stage was used in order to check whether the potential decrease in performance at S5-diff could not be attributed to fatigue only.

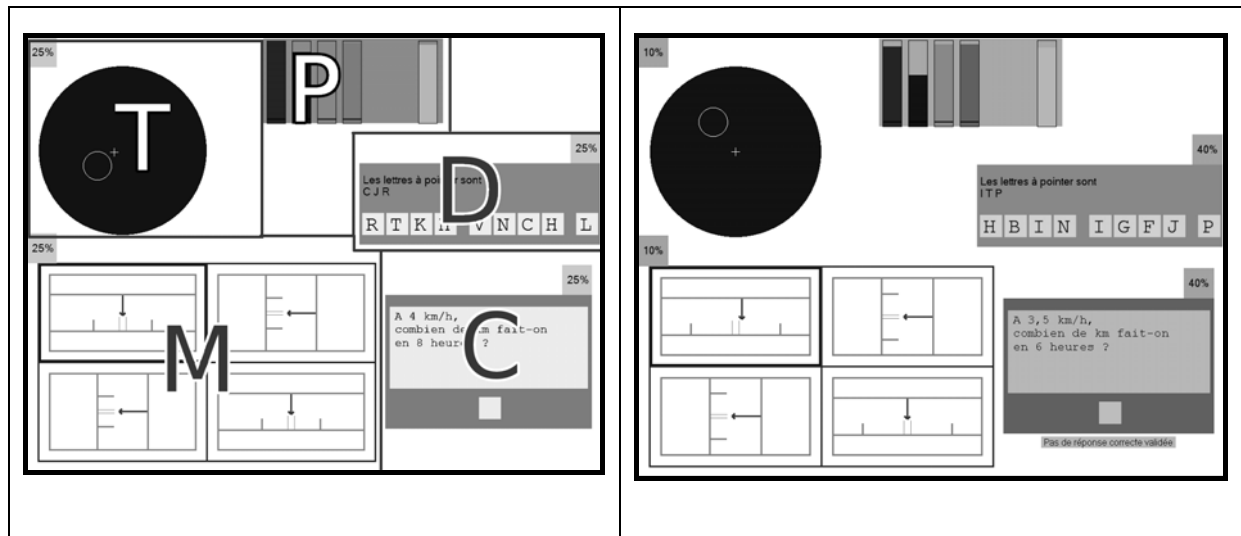


Figure 1. Screenshot of an equal-value stage, S4-equal or S6-equal (left panel) and of the differential-value stage, S5-diff (right panel). The letters on the left panel (*T* stands for tracking, *M* for monitoring, *D* for detecting and *C* for calculating subtask) have been superimposed here for the reader but were not displayed during the task. *P* stands for performance gauges. Percentages in grayed boxes represent the relative weight of each subtask for the computation of the global performance.

### Performance measurements

Performance indices were computed in 100ms steps. For each of the subtasks and for the global index as well, it ranged from 0 (worst) to 100 (best). Performance on Tracking was proportional to the distance between the cross and the moving circle. It was given a score of 100 whenever the cross remained within the moving circle. Depending on the speed of the moving circle, performance could decrease at a rate of 0 to 20% per second. Performance on Monitoring was proportional to the maximum distance between the four gauge levels and their corresponding target intervals. Moreover, the performance was given a score of zero whenever one gauge level went beyond a 60% tolerance interval. Depending on the speed of the gauge, performance could decrease at a rate of 0% to 33% per second. Performance on

Detecting and Calculating followed the same principle: Performance was given a score of 100 when the block of letters or the arithmetic problem was presented. Then the performance level gradually fell at a rate of 6.67 per second until the correct answer was keyed in. If a wrong answer was supplied, performance was more substantially decreased. Finally, the four subtask-performances were also continuously aggregated into a global performance index, which was based on a weighted mean of the four subtask performances. The weights corresponded to the values (percentages of importance) assigned to each subtask. However, in order to avoid complete neglect of a subtask and to increase the payoff of the utility maximization strategy, the global performance was set to zero all the while performance in at least one of the subtasks fell below a minimum threshold of 5%. This threshold rule was also applied at S5-diff and was explicitly communicated to the applicants. Visually, the threshold was materialized by a black horizontal line on each subtask performance gauge.

### **Preliminary Large Scale Field Study**

#### **Method**

**Participants.** 450 applicants took the Priority Management Test during the 2011 pilot selection session at ENAC (the French national civil airline pilot training organization). The PMT is used to assess the applicants' multitasking ability. They were 91% male, with a mean age of 21.6 yrs ( $SD=2.9$ ). Applicants were informed that, for research purposes, their performances could be analyzed anonymously. Most of them ( $n=320$ , 71%) were attending this pilot selection process for the first time. The others had already followed this selection process at a previous selection session but failed at some step (they could have failed on one of the five other assessed cognitive ability dimensions, or at the group exercises/interviews step or at the oral English examination). Thus, we could categorize the applicants in two groups following their prior experience with the Priority Management Test, (i) first-takers ( $n=320$ ) and (ii) retakers ( $n=130$ ). Most of retakers ( $n=123$ ) scored above the threshold when they took the PMT the first time, so they were already rather good performers on this task.

Moreover, it is known that applicants prepare themselves for such high stakes selection processes, therefore even first-takers have some practice with analogous tasks, but we had no means to gauge this experience.

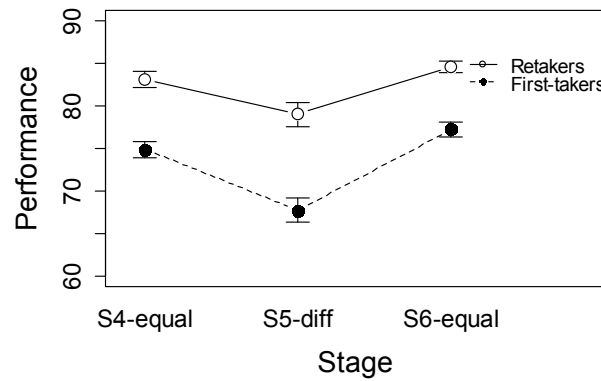
**Procedure.** Applicants sat in front of a desk with a computer screen, a keyboard and two joysticks. We used DELL 19" screens with a refresh rate of 75Hz and a resolution of 1024×768 pixels. Applicants were tested by groups of 20 to 30. The PMT was taken in the afternoon, after applicants had completed thirteen other cognitive ability tests.

**Analyses.** All analyses have been conducted with *R* software (R Core Team, 2014). Two factors Stage (3-levels within factor) x PriorExperience (2-levels between factor) mixed analyses of variance were conducted on global and subtask performance measurements. Tukey HSD tests were conducted to compare performances between S4-equal and S5-diff stages for first-takers and for retakers. The same post hoc test was used to compare S4-equal to S6-equal in order to test the potential impact of fatigue.

## Results

**Global performance.** Prior experience with the Priority Management Task had an impact on performance ( $F(1,448)=118.1$ ,  $p<.001$ ,  $\eta^2_p=.20$ ), with retakers performing 12% better than first-takers (see Figure 2). Moreover, the impact of the assigned priorities on performance depended on prior experience (significant Stage x PriorExperience interaction,  $F(2,896)=12.9$ ,  $p<.001$ ,  $\eta^2_p=.02$ ). Tukey HSD tests revealed a significant decrease in global performance for both first-takers and retakers ( $p<.001$  and  $p<.001$ ). However, the performance decrease between S4-equal and S5-diff was significantly greater for first-takers ( $M=-7.2$ ,  $SD=9.9$ ) than for retakers ( $M=-4.1$ ,  $SD=5.2$ ),  $t(448)=3.3$ ,  $p<.001$ ,  $d=0.35$ . Thus, first-takers seemed to have more difficulties than retakers to maximize global utility at the differential-value stage. Moreover, mean performance at S6-equal was not inferior to mean

performance at S4-equal (no significant difference for retakers,  $p=.31$ , and a significant increase for first-takers,  $p<.001$ , with the Tukey HSD test).



*Figure 2.* Mean global performance (with 95% confidence intervals) of the two groups identified by their prior experience with the Priority Management Task (retakers and first-takers) at the three four-subtasks stages (S4-equal, S5-diff and S6-equal).

## Discussion

The results of this preliminary large scale field study highlighted differences in performances between participants having prior experience with the multitasking environment and those having no known experience of such environment. Consistently with our hypothesis, at the differential-value stage first-takers had lower global performances than retakers. Moreover, the lack of decrease in global performance at the last stage, ruled out the argument that the decrease at S5-diff could be attributed to fatigue. Thus, the decrease in performance can be attributed to the change in subtask weights, as it was the only element that had changed. In addition, the increase in global performance at S6-equal for the less experienced participants was consistent with Gopher's (2007) findings on benefits of emphasis change on learning. Indeed, the emphasis change, although leading to detrimental effects on performance during the period of change, led to an increase in performance when all weights became equal again.

However, in order to more precisely characterize differences in resource reallocation strategies between the equal-value and differential-value stages, we conducted a laboratory study including eye movement recording.

### **Laboratory eye-tracking Study**

#### **Hypothesis**

We hypothesized that at the differential-value stage, novices would follow a locally rational strategy, whereas more experienced participants would follow a globally rational strategy. More specifically, the following hypotheses were tested.

*H1:* as well as in the preliminary study, global performance should be inferior at the differential-value stage than at the equal-value stage. Moreover, this difference in global performance should be greater for Novices than for Experienced participants.

*H2:* the decrease in global performance should be related to a decrease in performance at the two low-value subtasks, and this latter decrease should be greater for Novices than for Experienced.

*H3:* Novices, following a locally rational strategy should fixate less frequently the low-value subtasks and more frequently the high-value subtasks, at S5-diff than at S4-equal.

*H4:* Experienced, following a globally rational strategy should not fixate less frequently the low-value subtasks at S5-diff than at S4-equal. However, in order to maximize their payoff, Experienced should fixate the low-value subtasks when their performances come close to the minimum threshold (i.e. when it becomes urgent to process them).

#### **Method**

**Participants.** Thirty five participants performed the Priority Management task. In order to observe the effects of experience, we contrasted twenty pilot students (who had been selected based on their performance at this test) and 15 university students (novices regarding the task). Pilot students were mostly male (80%) and all aged between 19 and 24 ( $M=21.4$  yrs).

University students were 53% male and all aged between 22 and 29 years ( $M=25.6$  yrs). To estimate the amount of training hours devoted to the Priority Management Task, we asked the pilot students after they finished pilot training (in order to obtain honest answers) to give the number of hours they spent practicing on this task before the selection session. Only ten pilot students answered this questionnaire, but the results were informative: they declared having spent  $M=21.2$  hours,  $SD=16.1$  (range from 4 to 50 hrs). As a result, pilot students could be considered more experienced with this task than university students. Informed consent was obtained from each participant.

**Eye tracking.** Eye movements were recorded using an EyeLink 1000 desktop eye tracker (SR Research Ltd., Mississauga, Ontario, Canada). This eye-tracker possesses a spatial accuracy greater than  $0.5^\circ$  and a  $0.01^\circ$  spatial resolution. The sampling rate was set to 1000Hz. A chin and forehead rest was used to maintain these distances and to avoid head movements. All eye-tracking data were extracted using the SR Research default algorithm.

**Procedure.** The experiment took place in a quiet simulation room. The volunteers were seated on a comfortable chair. The eye tracker camera was placed at a distance of 20 cm from the screen and the eye camera was at a distance of 60 cm from the screen. We used the same computer screen as in the preliminary study. The lighting of the room was maintained constant. Before starting the experimental phase, participants performed a short calibration phase in order to adjust the eye tracker. Participants were then told to perform the task in strict accordance with the instructions provided on the screen.

**Analysis.** We conducted mixed-design analyses of variance (ANOVA) on performance and eye tracking variables, with Group as between-subjects factor (Group: Novices and Experienced) and Stage as within-subjects factor (Stage: S4-equal, S5-diff and S6-equal). Tukey HSD tests were used as post hoc tests in order to compare S4-equal and S5-diff as well as S4-equal and S6-equal. In addition,  $t$  tests and associated effect sizes (Cohen's  $d$ ) were computed to test specifically for changes in measurements across S4 and S5. When variances

differed greatly among groups, the Welch  $t$  test was computed instead. Concerning subtask variables, in order to test whether the type of subtask had an impact on the dependent variables we also conducted a three-factor (Subtask x 3 Stage x 2 Experience) mixed ANOVA. For eye tracking recordings we defined four areas of interest (AOIs), one for each subtask and computed dwell times spent in each AOI, that is the cumulative time spent fixating a given area. Performance and eye tracking data were also analyzed at a more detailed level in order to explore the dynamics of the strategies used by Novices and Experienced participants. Indeed, as an event was triggered on each subtask every 15s at the same time, we analyzed the evolution of performance and eye tracking variables for each subtask in 15s slots. For each of the sixteen 15s slots of the 4 min duration of a stage, we averaged data by blocks of 3s. Consequently we could follow the evolution of the variable after the trigger of an event that required an action from the participant.

## Results

**Global Performance.** Global performances (see Figure 3, top panel) were significantly greater for Experienced than for Novices,  $F(1,33)=87.2$ ,  $p<.001$ ,  $\eta^2_p=.72$ . Moreover, Novices' performance dropped ( $p<.001$ ) after task priority switched from S4-equal to S5-diff but Experienced participants' performance did not change significantly ( $p=.75$ ). Indeed, the Experience x Stage interaction was significant,  $F(2,66)=12.0$ ,  $p<.001$ ,  $\eta^2_p=.26$ . In addition, the decrease in global performance was significantly greater for Novices ( $M=-18.3$ ,  $SD=21.0$ ) than for Experienced participants ( $M=-3.6$ ,  $SD=4.4$ ),  $t(14.9)=2.65$ ,  $p=.02$ ,  $d=1.04$  (as variances differed by a factor of almost five across groups, we used Welch's  $t$  test). Thus, our results supported *H1*. Moreover, global performance at S6-equal again reached a similar level to S4-equal ( $p=.99$  for Experienced participants and  $p=.56$  for Novices), ruling out the fatigue hypothesis for the decrease in performance at S5-diff.

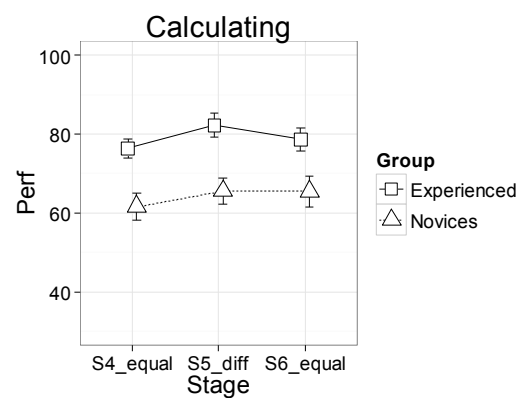
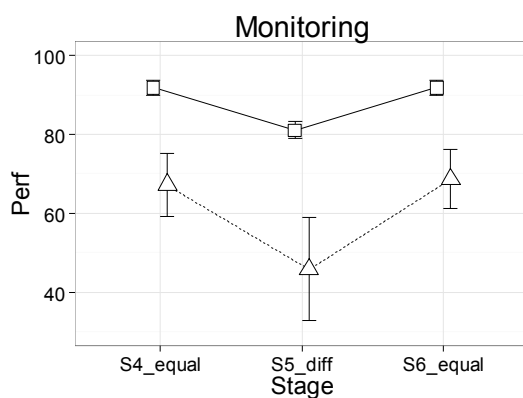
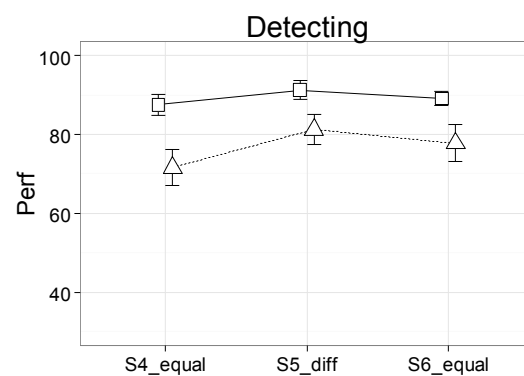
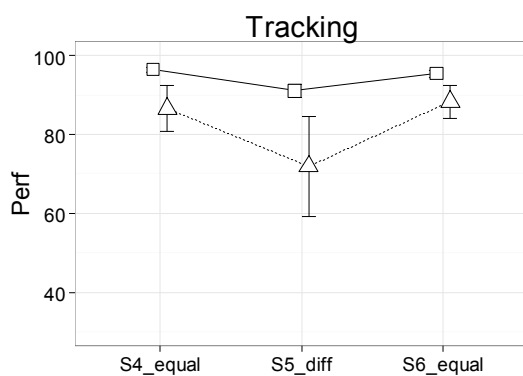
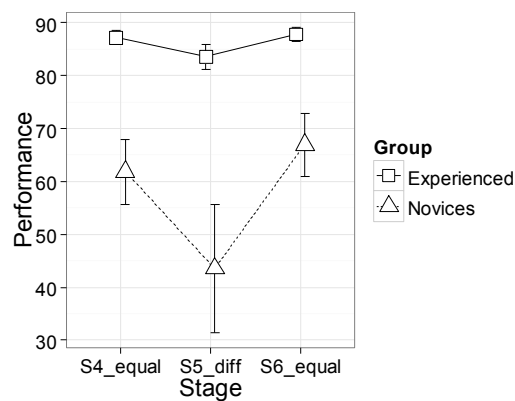
**Subtask Performances at the overall stage level.** The 4 Subtask x 3 Stage x 2 Experience mixed ANOVA showed a significant three-order interaction (see Figure 3);



$F(6,198)=2.85$ ,  $p=.01$ ,  $\eta^2_p=.08$ . For Novices, Tukey HSD tests between performances at S4-equal and S5-diff stages revealed (i) significant decreases for both low-value subtasks ( $p<.001$  and  $p<.001$ ), (ii) a significant increase for one high-value subtask ( $p=.046$ , for Detecting), and (iii) no significant difference for the other high-value subtask ( $p=.99$  for Calculating). The same post hoc tests for Experienced participants revealed (i) a significant decrease on only one low-value subtask ( $p<.001$  for Monitoring) and (ii) no significant difference for the other subtasks ( $p=.76$ ,  $.99$  and  $.62$  for Tracking, Detecting and Calculating). To sum up, after the emphasis on two subtasks, Novices had poorer performances on both low-value subtasks whereas Experienced participants had poorer performance only on one low-value subtask (Monitoring). However, for the two low-value subtasks, the performance decreases were not significantly greater for Novices than for Experienced participants while corresponding effect sizes were medium:  $p=.11$ ,  $d=0.65$  for Monitoring ( $M=-21.2$ ,  $sd=23.8$  for Novices and  $M=-10.8$ ,  $sd=5.4$  for Experienced participants) and  $p=.12$ ,  $d=0.64$  for Tracking ( $M=-14.6$ ,  $sd=21.7$  for Novices and  $M=-5.4$ ,  $sd=3.0$  for Experienced participants). Therefore our results only partially supported  $H2$ : The decrease in global performance was related to a decrease in performance at both low-value subtasks for Novices but only at one low-value subtask for Experienced participants. Moreover, this decrease was not significantly higher for Novices than for Experienced participants. Additional analyses were conducted in order to interpret this lack of significant difference for the two low-value subtasks. Specifically, we computed Fisher tests to compare the variances of the change in performance between S4-equal and S5-diff for both groups. The variance of the change in performance for both low-value subtasks was significantly greater for Novices than for Experienced participants,  $F(19,14)=0.05$ ,  $p<.001$  for Monitoring and  $F(19,14)=0.02$ ,  $p<.001$  for Calculating. Thus, the lack of significant difference of change in performance between Novices and Experienced participants was probably related to the large variability in performances at S5-diff for Novices for both low-value subtasks. Indeed Novices

## ALLOCATION STRATEGIES AFTER SWITCH IN VALUES

performances ranged from 2.0 to 78.4 for Monitoring and from 8.1 to 96.1 for Tracking. Thus, the way Experienced participants adapted to the change in emphasis at S5-diff was homogeneous whereas this change led to various effects for Novices, especially for both low-value subtasks. The next section will explore at a more detailed level the way participants adapted their resource sharing after the change in emphasis.



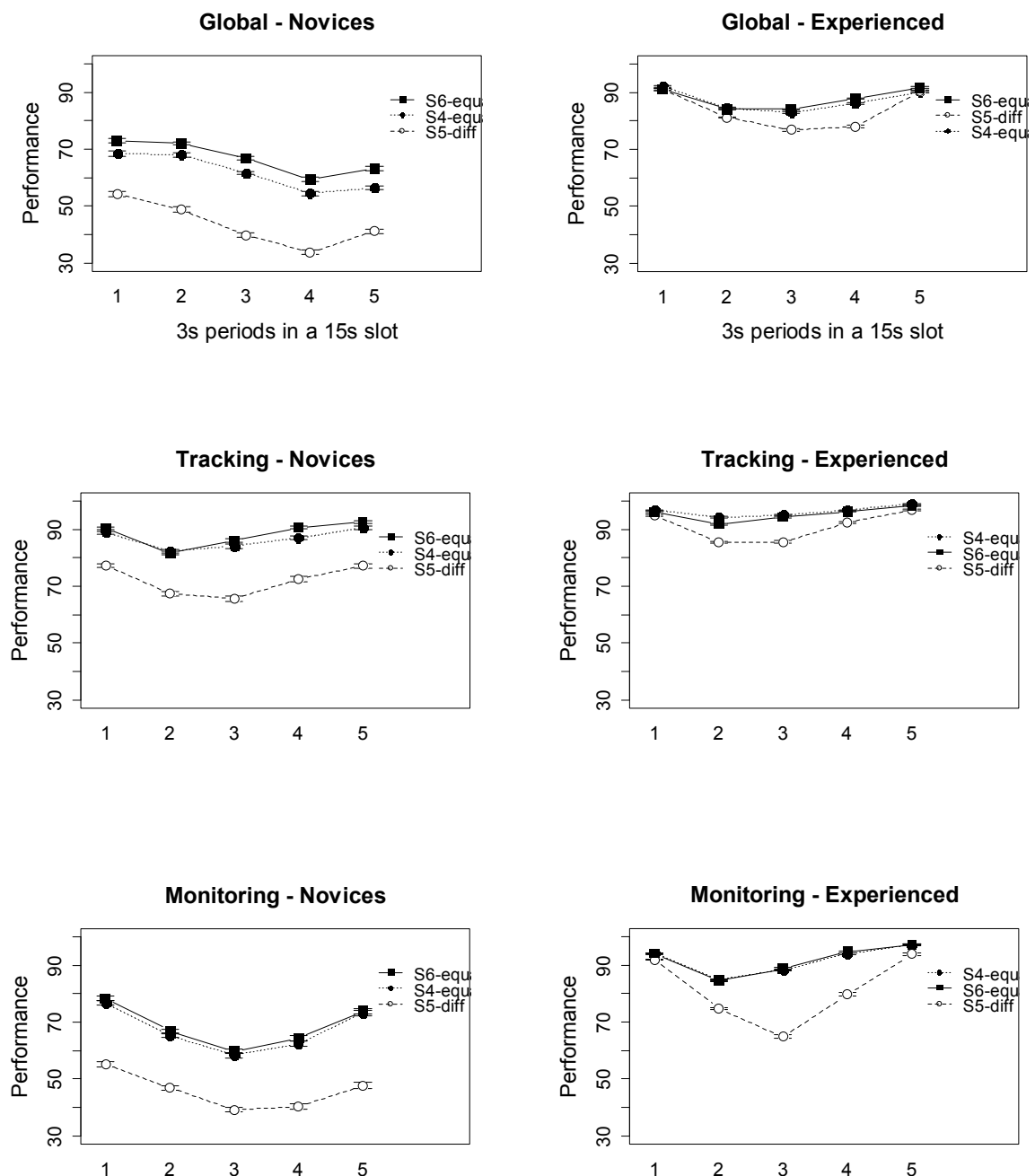
*Figure 3:* Mean global and subtask performances (0-100) for Novices (dashed lines) and Experienced participants (solid line) at each stage, with 95% confidence intervals (some CI are not missing but too small to be visible).

**Global and subtask Performances at a fine grained level.** Performance was analyzed in 15s slots, which correspond to the rate of events triggering for each subtask. For Novices (see Figure 4, left panel), global performance at S5-diff was inferior to that at S4-equal on the whole 15s slot (all Tukey post hoc tests  $ps < .001$ ). In fact this inferior performance was observed for the two low-value subtasks in the whole 15s slot (all Tukey post hoc tests'  $ps < .001$ ). On the contrary, performances on the high-value subtasks were superior when these subtasks were emphasized. One high-value subtask, Detecting, was better performed from the middle of the 15s slot (Tukey post hoc tests'  $ps < .001$  from the third period of 3s). For the other high-value subtask, Calculating, the increase in performance was only significantly observed at the end of the 15s slot ( $p < .001$  at the fifth period of 3s). Indeed, computing a mental calculation is more time consuming than identifying a target among distractors. In summary, Novices appeared to adapt to the emphasis change by processing the high-value subtasks earlier and by performing poorly on the two low-value subtasks from the beginning of the 15s period. This poor performance on both low-value subtasks was associated with a higher proportion of time spent below the minimum performance threshold (see Figure 5) at the differential-value stage compared to the equal-value stage.

Concerning Experienced participants (see Figure 4, right panel), the change in emphasis at S5-diff led to visible poorer performances at both low-value subtasks after 3s ( $ps < .001$  at periods 2 and 3 for Tracking and at periods 2, 3 and 4 for Monitoring) compared to S4-equal. At the same time, they had better performances on one high-value subtask, Calculating, after 6s ( $ps < .001$  for periods 3 to 5). However, on the other high-value subtask (Detecting) performance was not significantly better at any period (Tukey's  $ps$  ranged from .11 to .99),

## ALLOCATION STRATEGIES AFTER SWITCH IN VALUES

which could be due to a ceiling effect. However, the performance for both low-value subtasks reached the same level at S5-diff stage as at S4-equal at the end of the 15s slot ( $p=.99$  at period 5 for Monitoring and Tracking).



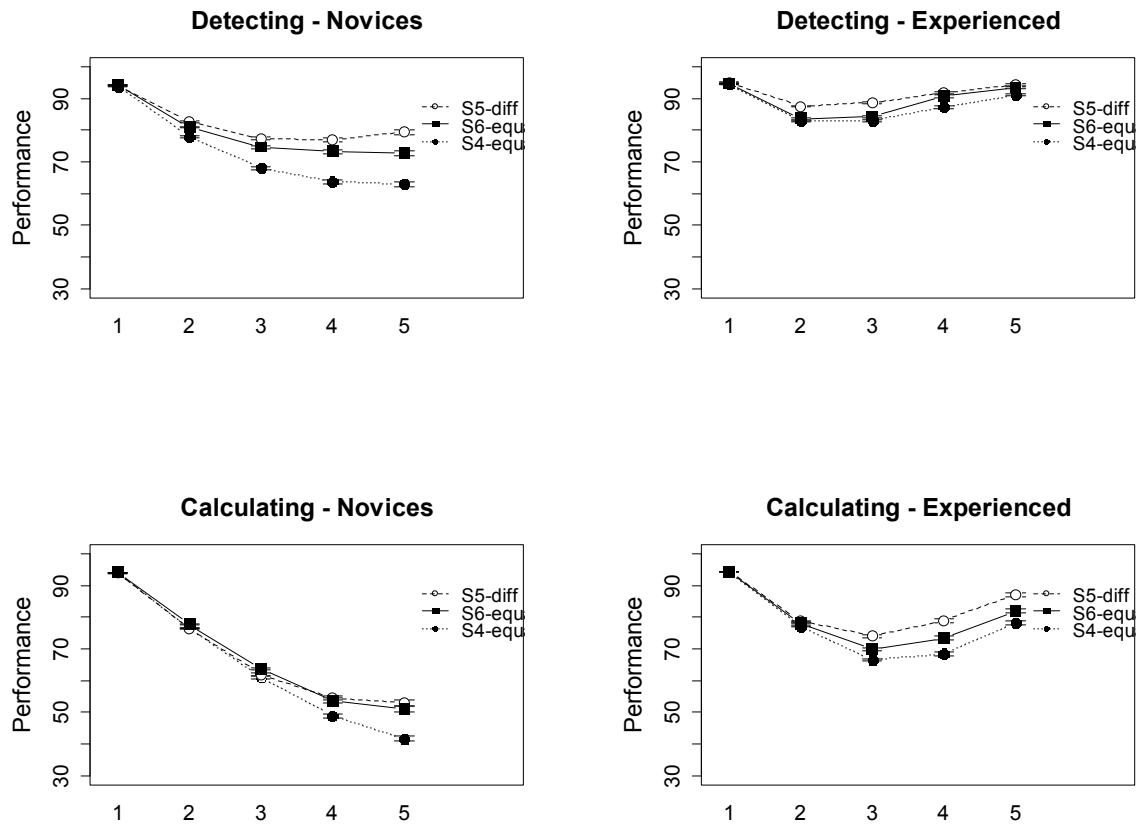


Figure 4. Mean global and subtask performances (0-100) with 95% confidence intervals for each phase aggregated over 3s periods during a 15s slot, and over the 16 slots of each 4mn stage, for Novices (left panel) and Experienced participants (right panel).

Moreover, the proportion of time spent below the minimum performance threshold for the two low-value subtasks was similar at S5-diff and at S4-equal (see Figure 5, right panel). Indeed, the only significant difference was observed for the fourth 3s period of the Monitoring subtask ( $p < .001$ ), but Experienced participants still spent a low proportion of time below threshold (8%). More generally, the absolute value of the proportion of time spent below the threshold was low (10%) compared to Novices (30 to 40%). To sum up, Experienced participants seemed to adapt to the emphasis change by processing earlier at least one high-value subtask, while preventing performances of the low-value subtasks from falling below the minimum threshold. These results needed to be confirmed with the eye

tracking data analysis.

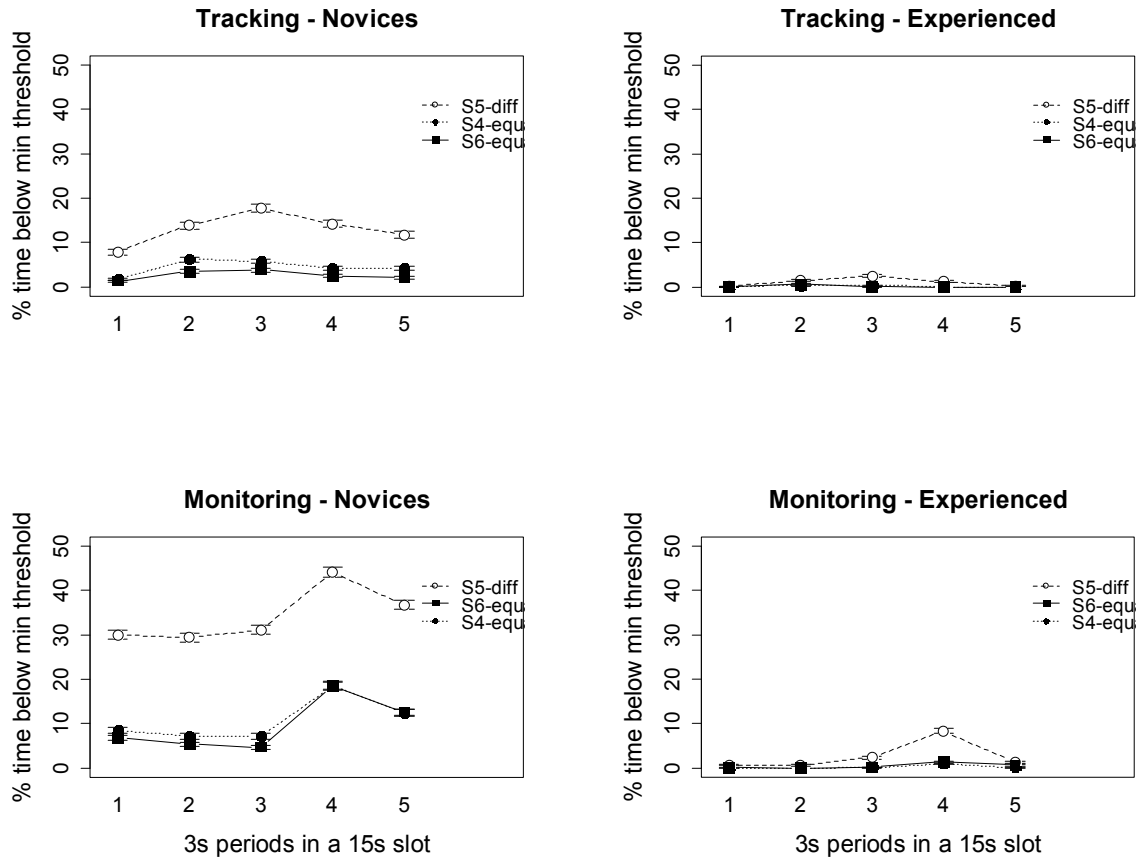
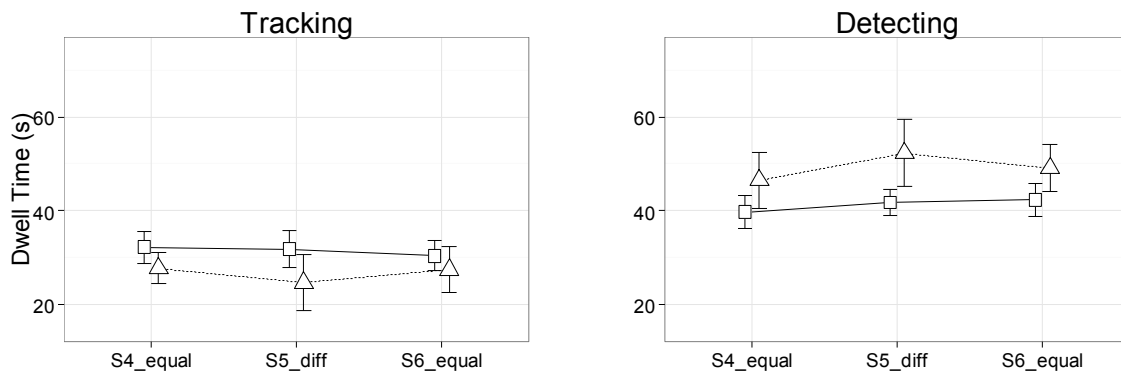


Figure 5. Proportion of time spent below minimum threshold for the two low-value subtasks with 95% confidence intervals for each phase aggregated over 3s periods during a 15s slot, and over the 16 slots of each 4mn stage, for Novices (left panel) and Experienced participants (right panel).

**Subtask attention reallocation strategies at the overall stage level.** The dwell times on each subtask highlighted that Experienced participants and Novices shared their visual attention differently among the four subtasks and reacted differently to the priority change (cf. Figure 6). Indeed, the third order Stage x Group x Subtask interaction was significant for dwell times,  $F(6,198)=3.13$ ,  $p=.006$ ,  $\eta_p^2=.09$ . For Experienced participants, Tukey HSD tests between the S4-equal and S5-diff showed no significant difference in dwell times for any subtask ( $p=.99$  for each subtask), thus supporting *H4*. However, for Novices, dwell times decreased significantly ( $p<.001$ ) for one low-value subtask (Monitoring). The differences in dwell times were not significant for the three other subtasks ( $p=.99$ ,  $.99$  and  $.93$  for Tracking, Detecting and Calculating). Accordingly, our results only partially supported *H3*, as we expected Novices to also fixate less frequently the other low-value subtask and to fixate more frequently both high-value subtasks.



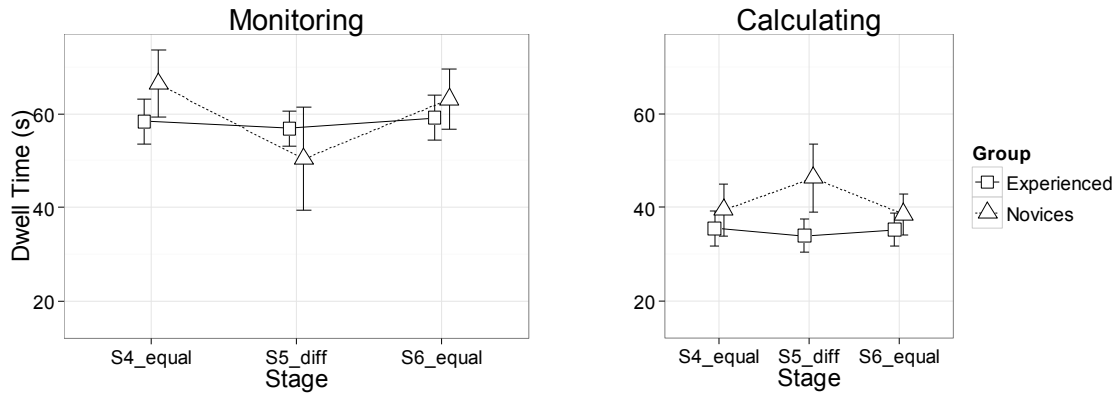


Figure 6. Mean dwell times in seconds and 95% confidence intervals for the four subtasks and the two groups (solid line for Experienced participants and dashed for Novices) at each stage.

Firstly, one could question why the decrease in dwell time was not significant for the second low-value subtask (Tracking) for Novices. One interpretation could be related to the low cognitive effort required by this subtask compared to the other subtasks, and to a ceiling effect. Indeed, in both equal-value stages, among the four subtasks, the Tracking subtask received the lowest dwell time ( $M=27.7s$  for Novices and  $M=32.2s$  for Experienced participants). Therefore, the potential decrease in dwell time could only be smaller. Moreover, performance data revealed a significant decrease in Tracking performance at S5-diff for Novices. Thus, a lack of significant decrease in dwell times was not related to a lack of significant differences in performance. Secondly, the lack of differences in dwell times between S4-equal and S5-diff for Novices was associated to an increase in performance for Detecting and to a lack of difference in performance for Calculating. One could suppose that it was easier to increase one's performance at detecting more target letters than at finding the correct answer to the mental calculation within a period of 15s. Finally, one could question the lack of differences in attention allocation for Experienced which could be interpreted as a lack of change in resource allocation strategy. However, this lack of difference was observed for total dwell times recorded over each whole 4-mn phase. Thus, a more detailed analysis



was required in order to explore the attention allocation strategies during the 15s periods following an event on each of the four subtasks (see next subsection).

**Fine-grained temporal reallocation strategies.** To evaluate whether Experienced participants changed the order in which they processed the four subtasks after an event was triggered on each subtask at the same moment (every 15s), we analyzed the evolution of gaze durations for each subtask during these 15s slots in five blocks of 3s. Experienced participants changed their allocation strategies when the priorities changed for one of the four subtasks (see Figure 7). Indeed at S5-diff, Experienced participants devoted less attention to one low-value subtask (Monitoring) at the beginning of the 15s slot (Tukey's HSD  $p < .001$  for first and second periods of 3s) and devoted a greater attention to this subtask later ( $p < .001$  at the fourth period of 3s). All other differences in dwell times between S4-equal and S5-equal were non significant. In other words, at S5-diff Experienced participants changed their attention allocation strategy by delaying the resource allocation of the most demanding low-value subtask in the 15s slot after the trigger of a task event. However, this was not associated with a significantly greater attention for the high-value subtasks.

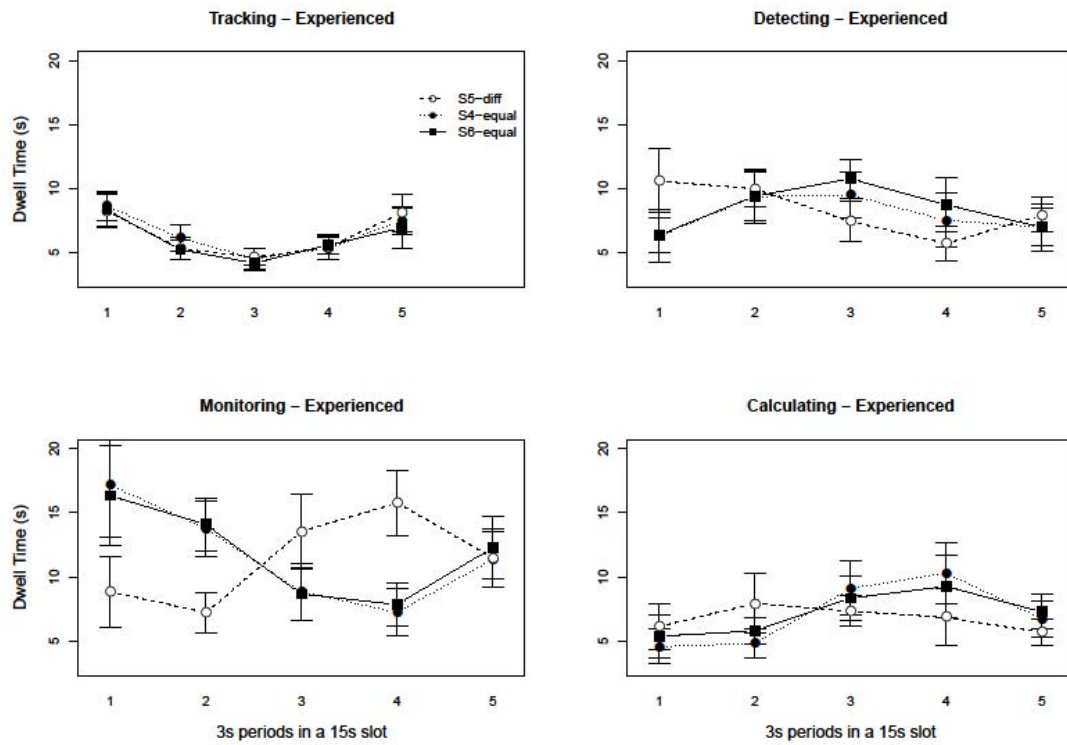


Figure 7. Mean dwell times (in seconds) and 95 confidence intervals for each subtask and each phase aggregated over 3s periods during a 15s slot, and over the 16 slots of each 4mn stage, for Experienced participants. At S5-diff, Detecting and Calculating were the high-value subtasks and Tracking and Monitoring were the low-value subtasks.

On the results for Novices the same analyses showed (see Figure 8) that they devoted significantly less attention to the same low-value subtask at S5-diff than Experienced participants, i.e. Monitoring, shortly after the triggering of the task event ( $p=.04$  at the second period of 3s). All other differences in dwell times between S4-equal and S5-equal were non significant. Accordingly, unlike Experienced participants, Novices did not reallocate attention to Monitoring at the end of the 15s period.

## ALLOCATION STRATEGIES AFTER SWITCH IN VALUES

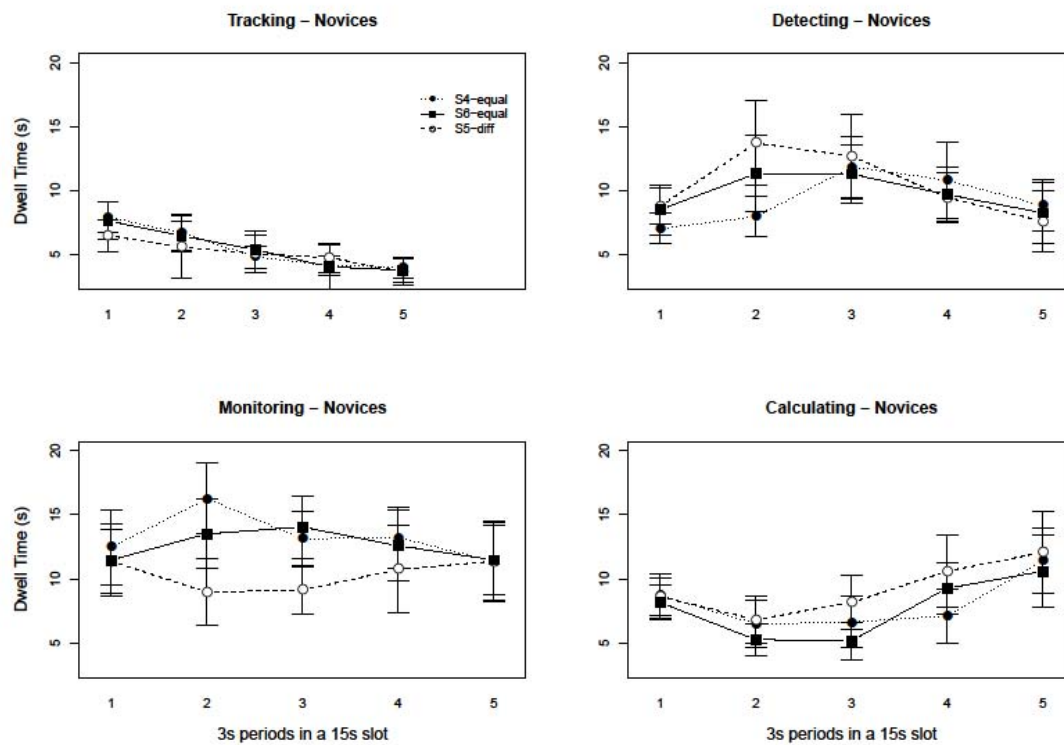


Figure 8. Mean dwell times (in seconds) and 95% confidence intervals for each subtask and each phase aggregated over 3s periods during a 15s slot, and over the 16 slots of each 4mn stage, for Novices. At the S5-diff, Detecting and Calculating were the high-value subtasks and Tracking and Monitoring were the low-value subtasks.

### Discussion

Several multitasking models describe behavior when dividing attention among several subtasks with different values. For instance, the STOM model predicts that people will devote more time to a high-value subtask than on a low-value subtask. We investigated individual differences in the management of a change in subtask values in a multitasking environment. Specifically we hypothesized that novices would follow a locally rational strategy and devote less attention to the low-value subtasks which would lead to lower global performances in the differential-values stage than in the equal-values stage (*H1*). On the contrary, experienced participants were expected to follow a globally rational strategy and devote their attention to the low-value subtasks only when it becomes an imminent threat for global performance (*H2*).

*H1* was supported by our results as global performances were inferior at the differential-value stage than at the equal-value stage and this decrease was greater for Novices than for experienced participants. In other words, Novices had better global performances when they divided their attention among four subtasks than when they were instructed to focus on two subtasks without ignoring the two others. *H2* was only partially supported. Indeed, Novices performance decreased for both low-value subtasks whereas experienced participants performance decreased only for one low-value subtask. Thus, results were consistent with Novices following a locally rational strategy, as they did not prevent the two low-value subtask performances from falling below their minimum threshold. However, the fact that experienced participant performance decreased for only one low-value subtask could be interpreted in line with the globally rational strategy. Indeed, experienced participants were likely to have enough cognitive resources to maintain a similar performance at least on one low-value subtask. Moreover, our findings highlighted the large variability in behavioral response of Novices, especially for the two low-value subtasks at the differential-values stage.

*H3* and *H4* concerned eye tracking data. *H3* was only partially supported, as Novices fixated significantly less only one low-value subtask (Monitoring). Nevertheless, for the other low-value subtask (Tracking), performance decreased at the differential-value stage. Therefore one interpretation could be that Novices devoted the same amount of visual attention to this subtask, but did not relate this perception to the same amount of actions. *H4* was supported by our results, so eye tracking data confirmed the globally rational strategy of Experienced participants. In other words, Experienced participants devoted (1) either the same amount of visual attention at the differential-values and at the equal-values stages or (2) less visual attention to a low-value subtask until it became a threat for the global performance.

To sum up, Novices adapted their resource allocation strategy by putting less emphasis on the most demanding low-value subtask and putting more emphasis on the easiest high-value subtask. Thus, Novices reacted to the change in emphasis by a strategy that favored one local utility but had detrimental effects on the global utility. On the contrary, Experienced adapted to the change in emphasis by devoting the same attention for both high-value subtasks which guaranteed satisfying levels of performance. However, on both low-value subtasks, Experienced participants still devoted enough attention in order to prevent their performances from falling below the minimum threshold. Therefore, Experienced participants still maximized the global utility at the differential-value stage.

### **General Discussion**

The present findings showed that in a multitasking context with a high rate of events to process and an emphasis on some subtasks, people differently reallocated their attention among high-value and low-value subtasks depending on their previous experience with this multitasking environment. Novices devoted more attention to one high-value subtask and less attention to the more demanding low-value subtask which had detrimental effects on the

global performance. On the contrary, experienced participants still devoted attention to this low-value subtask especially when it became a threat for the global performance. Thus, consistently with our hypothesis, novices followed a locally rational strategy whereas experienced participants followed a globally rational strategy. These results extend those of Katidioti and Taatgen (2014) by taking into account individual differences in the degree of familiarity with the multitasking environment. These findings suggest that participants who are familiar with a multitasking environment may have sufficient available resources to follow a global rationality principle, whereas participants who are new to the task have fewer available resources and will follow a local rationality principle. This local rationality principle led novices to discard the most demanding low-value subtask. Indeed, neglecting low importance subtasks has been previously identified as one of the basic strategies to resolve a resource conflict (Freed, 1998; Schneider & Detweiler, 1988).

Threaded cognition (Salvucci & Taatgen, 2011), with its urgency principle, described the behavior of experienced individuals. Indeed, they processed the low-value subtask as soon as it became urgent to correct the deviation from target, which prevented a dramatic decrease of global performance. On the contrary, novices apparently did not follow this urgency principle, as the performance of both low-value subtasks remained a substantive amount of time below minimum threshold, which led to a dramatic decrease of global performance. Novices adhered to a minimum control principle (e.g., Taatgen, 2007) by preferring optimizing performance on two rather than four subtasks. Moreover, among the two high-value subtasks, novices favored the easiest one, in accordance with the effort avoidance principle (e.g., Kahneman, 2011). The STOM model (Wickens & Gutzwiller, 2015) would have predicted a preference for the high-value subtasks. Indeed, in a MATB-II (Multi-Attribute Task Battery) environment experiment, Gutzwiller (2014) found that participants spent more time on the high-value subtask when this subtask was more difficult. He found no effect of subtask value when this subtask was easy. Our findings highlighted that the preference for high-value

subtasks may be modulated by task experience. Indeed, our results suggested that novices, for whom the whole multitasking environment is difficult to manage, would have the tendency to take task value more strongly into account than individuals who are more familiar with the multitask environment.

Moreover, our results from novices are consistent with previous findings for untrained participants. In another multitasking environment, SYNWORK1, with variation in emphasis on subtasks, participants modified their strategies when subtask payoffs changed (Wang, Proctor & Pick, 2007). Specifically, Wang et al. (2007) observed that some participants almost abandoned a subtask after the decrease in this subtask emphasis (a memory task). Similarly, in the SynWin multitasking environment, participants tended to focus on one subtask out of four in an emergency condition whereas they alternated between several subtasks in the baseline condition (Hambrick, Oswald, Darowski, Rench & Brou, 2010).

### **Limitations**

The task used in this study has the advantage of being ecologically valid, but has several limitations concerning the generalization of our findings. Firstly, the four subtasks require very different cognitive processes. We observed the effects of the emphasis on two subtasks, but further research would be needed to test the effects of the emphasis on two other subtasks. Secondly, further research would be needed to test the effects of the emphasis on only one subtask. Thirdly, no condition tested the alternate prioritization scheme, i.e. starting by differential values first. Finally, we contrasted two very different kinds of populations, participants who were totally new to the task and not particularly selected on the basis of their cognitive abilities, and those who were very familiar with the task but also selected for their general high cognitive abilities. Thus, further research would be needed in order to disentangle the effect of familiarity with the task and general cognitive ability.

### **Practical Implications**

In a context of complex skill learning where trainees have to learn to manage several tasks simultaneously, the variable priority paradigm has proven to be useful (see Gopher, 2007 for a review). However, present findings showed that a majority of novices, following a locally rational strategy, would have the tendency to discard a subtask explicitly labelled as less important than the others. If the task requires not completely abandoning any subtask, instructors could for example offer part-task training first, in order to reduce the cognitive load of the trainee for each subtask. Then, during the whole-task training instructors could use both equal-priority and differential-priority stages. Indeed, equal-priority stages force trainees to divide their attention among all subtasks, and to find strategies that enable a satisfying global performance to be obtained. Finally, in order to promote metacognitive knowledge about the task, instructors could explicitly warn trainees about their tendency to neglect low-value subtasks. All these recommendations require further research to be validated. Moreover, given the variability of novices performance at the differential-priority stage, one cannot suppose that only one training strategy suits all trainees.



### Key Points

- We focused on a high demanding multitasking environment where subtask value vary.
- After the change in emphasis, performance dropped more for novices than for experienced participants
- Novices tended to follow a resource saving strategy whereas Experienced participants follow a utility maximization strategy.

### References

- Bellenkes, A.H., Wickens, C.D., Kramer, A.F. (1997). Visual Scanning and Pilot Expertise: the Role of Attentional Flexibility and Mental Model Development. *Aviation, Space and Environmental Medicine*, 68, 569-579.
- Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., ... Low, K. A. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta psychologica*, 135(3), 349-357.
- Duncan, J. (1990). Goal weighting and the choice of behaviour in a complex world. *Ergonomics*, 33(10-11), 1265-1279.
- Crundall, D. E., & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, 41, 448-458.
- Freed, M. (1998, July). Managing multiple tasks in complex, dynamic environments. In AAAI/IAAI (pp. 921-927).
- Fu, W.-T., & Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology*, 52, 195-242.
- Gopher, D. (2007). Emphasis change as a training protocol for high-demands tasks. *Attention: From theory to practice*, 101, 209-224.

- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, 71, 147-177.
- Gutzwiller, R. S. (2014). Switch choice in applied multi-task management. Doctoral Dissertation. Colorado State university.
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, 24, 1149-1167.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Katidioti, I., & Taatgen, N. A. (2014). Choice in multitasking how delays in the primary task turn a rational into an irrational multitasker. *Human Factors*, 56, 728-736, DOI: 10.1177/0018720813504216..
- Kramer, A. F., Larish, J. F., & Strayer, D. L. (1995) Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied*, 1, 50.
- Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2009). The multitasking myth. *Handling complexity in real-world operations*. Farnham: Ashgate.
- McLeod, P., Driver, J., Crisp, J. (1998). Visual search for a conjunction of movement and form is parallel. *Nature*, 332, 154-155.
- Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, 7(3), 134-140.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1-18). New York: Plenum.
- R Core Team (2014). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Robinski, M., & Stein, M. (2013). Tracking visual scanning techniques in training simulation for helicopter landing. *Journal of Eye Movement Research*, 6, 1-17.

- Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological review*, 115, 101.
- Salvucci, D. D., & Taatgen, N. A. (2011). *The multitasking mind*. Oxford University Press.
- Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling a connectionist/control architecture. *Human Factors*, 30, 539-566.
- Taatgen, N. A. (2007). The minimal control principle. In Gray W. (Ed.), *Integrated Models of Cognitive Systems* (pp. 368-379). New York: Oxford University Press.
- Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, 46, 629-646.
- Wang, D. Y. D., Proctor, R. W., & Pick, D. F. (2007). Acquisition and transfer of attention allocation strategies in a multiple-task work environment. *Human Factors*, 49, 995-1004.
- Wickens, C. D., & Gutzwiller, R. S. (2015). Discrete task switching in overload: A meta-analyses and a model. *International Journal of Human-Computer Studies*, 79, 79-84.
- Wickens, C. D., Gutzwiller, R. S., Vieane, A., Clegg, B. A., Sebok, A., & Janes, J. (2016). Time sharing between robotics and process control: Validating a model of attention switching. *Human Factors*, 58(2), 322-343.